

## CLAIMS:

1. A method of identifying the best matches or sets of matches between a query item and an item or items from a data set, comprising the steps of:
  - (i) providing a data representation for each item in the data set;
  - (ii) providing a query representation of the query item;
  - (iii) defining a transformation space;
  - (iv) for each of a number of regions spanning the entire transformation space, determining an upper bound to the probability of a match between the query representation and a data representation under any transformation in the region;
  - (v) determining a threshold probability;
  - (vi) comparing the upper probability bound of each region with the threshold probability; and
  - (vii) determining regions having an upper probability bound greater than the threshold probability, so as to identify solution regions.
2. A method as claimed in claim 1, and including the further steps of:
  - sub-dividing the solution regions into further regions which span the solution regions;
  - determining a new upper bound;
  - determining a new threshold probability; and
  - determining new solution regions.
3. A method as claimed in claim 2, including the step of iterating the further method steps of claim 2 so as to identify the solution region containing the best matching solution or to identify a set of solution regions containing a set of best matching solutions.

4. A method as claimed in claim 1, in which the data representations are topological representations of the data items and the query representation is a topological representation of the query item.
5. A method as claimed in claim 4, in which the topological representation of the data items and query item comprises a set of node measurement vectors, each node measurement vector being associated with a node of a topological arrangement of nodes defining the items.
6. A method as claimed in claim 1, in which the upper bound is determined using Bayesian probability theory.
7. A matching engine for identifying an item or items from a data set, the engine comprising electronic data processing apparatus including:
  - a memory storing a data representations for each item in the data set;
  - an input for inputting a query representation of the query item; and
  - a processor which includes means for defining a transformation space, means for generating a number of regions of the transformation space spanning the entire transformation space, means for determining for each region an upper bound to the probability of a match between the query representation and a data representation under any transformation in the region, means for determining a threshold probability, a comparison means which compares the upper probability bound for each region with the threshold probability, means to identify solution regions having an upper probability bound greater than the threshold probability,

and means to store an identification of a match between the query item and the item of the data set in a memory.

8. A computer program which when running on a computer carries out a method as claimed in claim 1.
9. Computer program code for identifying an item or items from a data set, the code including instructions for carrying out the functions of:
  - (i) providing a set of data representations of each item in the data set;
  - (ii) providing a query representation of the query item;
  - (iii) defining a transformation space;
  - (iii) for each of a number of regions of the transformation space spanning the transformation space, determining an upper bound to the probability of a match between the query representation and a data representation under any transformation in the region;
  - (iv) determining a threshold probability;
  - (v) comparing the upper probability bound of each region with the threshold probability; and
  - (vi) determining solution regions having an upper probability bound greater than the threshold probability, so as to identify the solution regions.
10. A computer readable medium storing computer code as claimed in claim 9.